Glare Reduction Using 4D Ray Sampling

Travis Allen
Stanford University
Palo Alto, CA
tallen1@stanford.edu

Abstract

For this project, I set out to reproduce the work done by Raskar et al. [6] back in 2008, only with a light field camera instead of a standard camera with pinhole array mask. I was able to replicate the basic method described in the original paper and went beyond this simple algorithm to explore some of the trade space in ray sampling. I looked at the effects of determining outliers using different color spaces as well as different percentages of rays per spatial tile (aka microlens). To evaluate these, I look at the color in the final images as well as relative luminance. I also present some statistical analysis of the ray space at each microlens and discuss how these methods could be used to develop a more sophisticated algorithm in the future.

1. Introduction

In traditional photography, image glare is a regular phenomenon. While glare can and has been used for artistic effect, for many practical reasons glare is actually quite undesirable as it can reduce overall image contrast and obscure image detail. Glare is been broadly characterized by Raskar et al. [6] as being either caused by reflection or caused by scattering. The errant light rays that are either scattered or reflected then saturate (either partially or fully) pixels that would otherwise capture the natural irradiance of a scene.

One can envision how glare, if not reduced or controlled to some extent, may negatively affect consumer photographs, sensors for autonomous vehicles, fluorescent microscopy, and astronomy. For many of these applications, deconvolution and other techniques are used in software to reduce the relative glare. On the physical side, lens designers use anti-reflective coatings and camera makers include baffles inside the camera body in order to significantly reduce the occurrences and effects of glare.

Glare has been traditionally viewed as a 2D problem, with many trying to characterize how glare manifested itself in 2D images. However, this particular project focuses on the reproduction of a method for reducing the effects of glare in an image using 4D ray sampling as described by Raskar et al. [6].

Specifically, this project makes use of a commercially available light field camera (the Lytro Illum). While the concept of light fields have been around for quite a while (since the early 1900s), the first commercial light field cameras weren't created until very recently. Unlike a standard digital camera, a light field camera focuses light not on the image sensor, but rather on an array of microlenses that is situated just in front of the sensor. One is then able to back out not only the irradiance at a particular point in a scene, but one can also get the direction from which each light ray is incident. Additional work accomplished in this area that was read for guidance included Levoy et al., Marwah et al., Ng et al, and Wetzstein et al. [2, 3, 5, 8]. Light fields have been used for myriad applications. For this project, a light field will be used to reduce glare in a manner that would be next to impossible with an unmodified, standard, commercial camera.

2. Related Work

As stated in the introduction, traditionally, glare has been treated as a 2D problem. Deconvolution methods were devised with glare Point Spread Functions (PSFs), as described by Bitlis et al. [1]. While somewhat successful, there are severe limitations that the method from Raskar et al. [6] actually helps address. Nayar et al. [4] described a method for separating direct and global components with high frequency illumination that required multiple exposures for high resolution results. While insightful, this method does is not practical for a handheld solution. Talvala et al. [7] explored methods for veiling glare in high dynamic range (HDR) images using a structured occlusion mask. This paper was the primary inspiration for the work that Raskar et al. [6] did in 2008. However, like some of the other proposals, it did not address single-shot, handheld applications.

Raskar et al. [6], building on the work from several of these other authors, proposed a method for reducing and characterizing glare using 4D ray sampling in 2008. At the

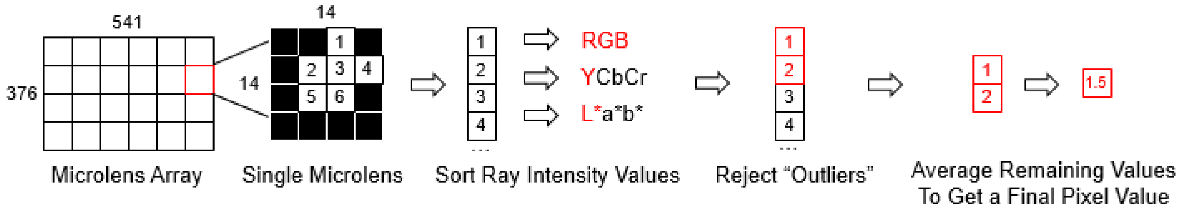

Figure 1. Flow Diagram for Basic Method

time, it was thought to be the first method that could significantly reduce glare in a single exposure (and I have seen no evidence to the contrary). While the paper discusses the use of light field arrays for the implementation of the glare reduction methods, they actually used a pinhole array mask placed in front of the sensor.

Due to the nature of the statistical ray sampling methods proposed in the paper, and the time at which the research was being done, the pinhole array masks made the most sense. The methods call for the reduction of spatial resolution in a light field array to the size of the microlens array (not the sensor size). Whereas, with a pinhole mask, one could get 4D information while retaining full sensor resolution. However, a pinhole mask also meant longer exposure times because the mask blocked a significant amount of light from the sensor and it also meant one needed to ensure the scene of interest was in focus as the out of focus areas would be heavily aliased. In a light field array, we don't have the same problem. We can also refocus after the fact using a shift since we aren't blocking out of focus rays from reaching the sensor. And, thankfully, since image sensors and manufacturing methods have come a long way in the past 10 years, we can get a fairly reasonable resolution from the Lytro Illum commercial light field camera. Thus, for this project, I will be reproducing and expanding upon the work done by Raskar et al. [6] in 2008. After some moderate searching, I could not find much that has been done in this area since this paper was published.

3. Method Employed

The basic method employed for this project was built upon the "Basic Algorithm" as described in section 3.1 of Raskar et al. [6]. The idea is fairly straightforward and is shown, in a very general way, in Figure 1.

In a traditional digital camera, an image is formed by the integration of the irradiance of all angular light rays incident at each photodiode on an image sensor. In a light field camera, we are focused on the microlens array instead of the image sensor. The rays pass through the microlens at various angles from the aperture and land on the underlying image sensor within a "tile". Thus, if we were to simply integrate over the irradiance values in the "tile" of pixels underneath each microlens, we would obtain a similar result to the 2D image formed by a traditional camera with an image resolution equal to the size of the microlens array. For the Lytro Illum (the camera used for this project), the image sensor resolution is about 40MP, but each microlens corresponds to a tile of about 14x14 pixels. The resolution of an image that does not do any additional post-processing besides simply averaging over each microlens is thus 0.2MP.

The basic method described by Raskar et al. (and that was used for this project), can be summarized in the following steps (also summarized in Figure 1):

1. Capture an image with the Lytro Illum portable light field camera.
2. Take the RAW image file and convert it to a demosaicked, rectified, light field image file with pixel values corresponding to angular values for each microlens (this turns out to be a 5264x7574x3 PNG image with "tiles" that are 14x14 for each microlens).
3. For each spatial sample of the light field image (i.e. each 14x14 "tile" or microlens), use some kind of statistical evaluation to determine outlier values among the angular samples.
4. Treat the outliers as glare and throw them out.
5. Average over the remaining values in each spatial sample and reconstruct a "low-resolution" 2D image.

Since we're looking at each spatial "tile" as a group of irradiance values regardless of the angle of incidence, the statistical operators one could use are fairly typical for any randomly distributed population. One could simply choose the median or the min ray irradiance in each and the glare would be significantly reduced. Of course, one of the primary benefits of having a light field camera is the fact that we could also sort based on incident angle. We could also refocus the image or look at the spatial frequency of the irradiance within each tile.

Raskar et al. suggested averaging over a percentage of the ray population in each tile as a starting point; specifically, over the darkest 20%. That was the starting point for

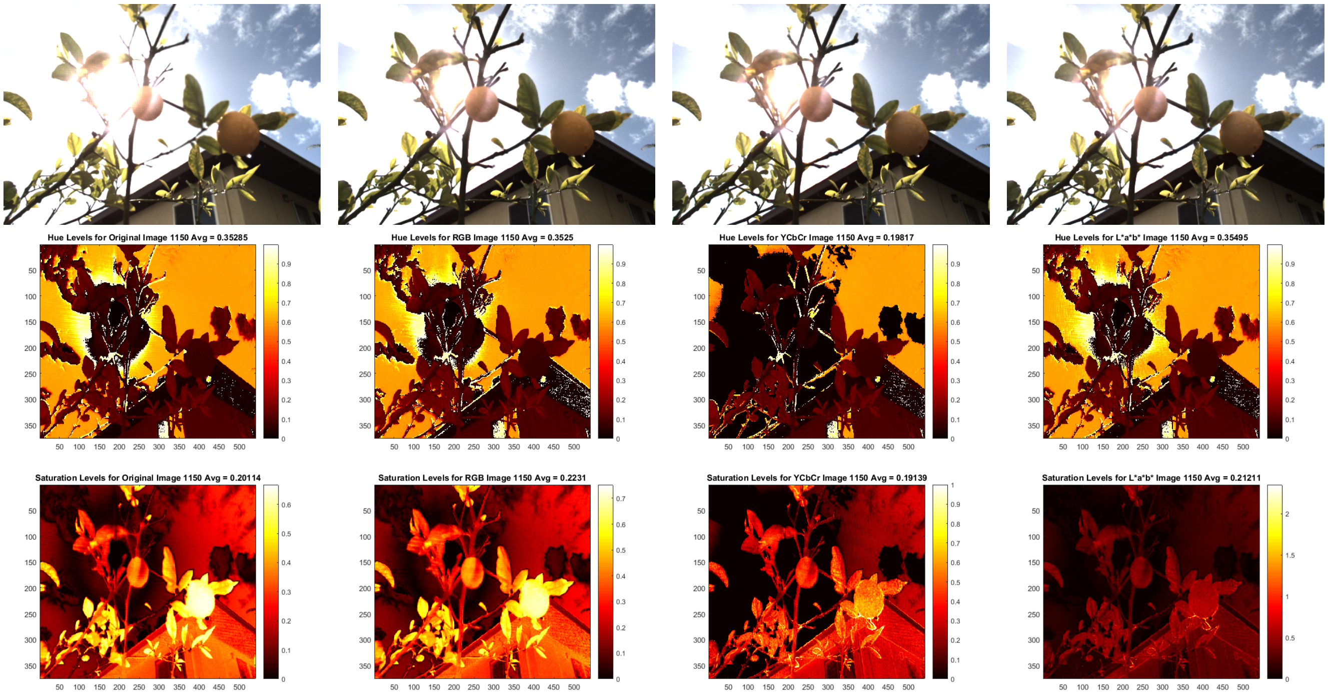

Figure 2. Comparison of Different Color Spaces for Outlier Determination [Top: Final Images / Mid: Hue Heat Maps / Bot: Saturation Heat Maps / L to R: Original Image (no glare reduction); RGB Manipulated; YCbCr Manipulated; L*a*b* Manipulated (all using 20% of angular samples)]

my particular method. I then looked at varying percentages of angular samples and the resulting effect on the luminance of an area of interest in the final image. I also wanted to look at whether or not it mattered what color space we used for the statistical determination of outliers. I looked at RGB (which was the native color space from the RAW image) as well as the luminance (Y) channel in YCbCr space and the lightness (L*) channel in CIE L*a*b* 1976 color space. Key here was to look at how these color spaces affected the color properties of the final image. I also looked at various statistical methods for exploring the ray space and how it could be manipulated to give better results.

One could actually conceptualize this process mathematically by considering the large array as a 4D array, $\mathrm{L}(x, y, u, v)$ with (x, y) being the coordinates for the image sensor plane and (u, v) being the coordinates in the aperture plane (ignoring color as a dimension/variable for the time being). Then the final 2D image will be $i(x, y)$ and can be given in a similar manner to equation 1. Simply taking the mean over (u, v) coordinates gives you the standard 2D image you would have gotten without removing outliers.

$$i_{orig}(x, y) = \mathrm{mean}_{u,v}[\mathrm{L}(x, y, u, v)] \qquad (1)$$

If we instead were to average over only the darkest 20% of ray values in each microlens tile, or even if we only took the minimum ray value, we would end up with an image with reduced glare, $i_{rg}(x, y)$. We could then get a pretty good picture of what the glare looks like by taking the maximum ray value in each tile and subtract the minimum, as in equation 2.

$$i_{glare}(x, y) = i_{max}(x, y) - i_{min}(x, y) \qquad (2)$$

One of the reasons this works so well is because, in many cases, the irradiance values of glare rays is fairly binary and one can easily separate glare values from the darker, more desirable values. As Raskar et al. [6] discuss in their paper, this method works best when the number of glare outliers is less than 50% of the spatial sample population.

4. Experimental Results

In order to explore how images could be manipulated from a raw 4D light field to a reduce glare 2D product, I decided to manipulate a few variables. The first was color space and the second was the percentage of rays that were not considered outliers. Finally, I looked at a few statistical operators in order to explore how glare affected the incident rays on the image sensor.

In this section I will present what I did along with figures and captions. In the next section (Analysis and Evaluation),

Figure 3. Comparison of Glare Reduction Based on Percentage of Rays Used as Outliers [L to R: Original Image (no glare reduction); 10% Rays Used; 25% Rays Used; 50% Rays Used (all manipulated in RGB)]

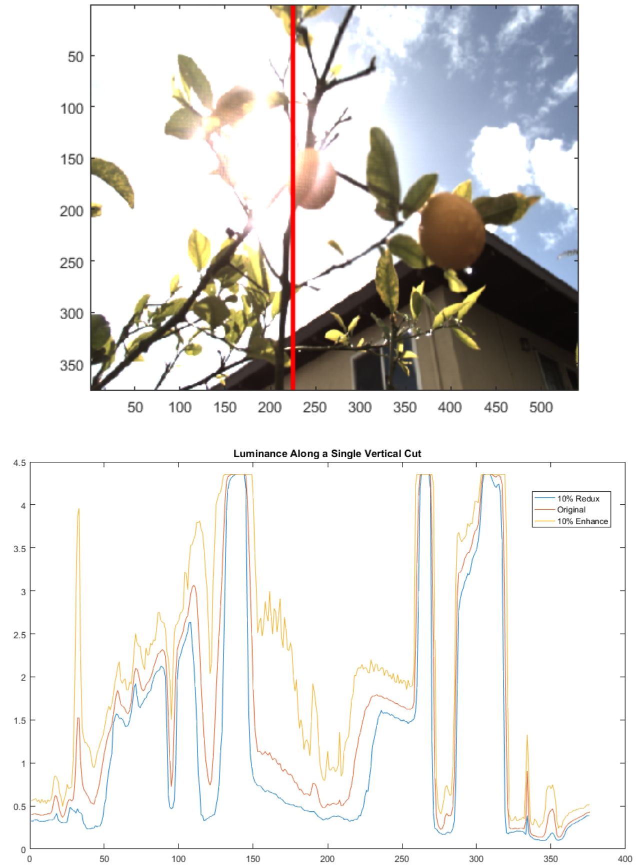

Figure 4. Luminance Along a Vertical Cut Through the Image of a Lemon – Graph Shows the Difference in Luminance for the Image Using an Average of the Lowest 10% of Rays for Glare Reduction, the Average Irradiance Values of All Rays for the "Original Image", and an Average of the Highest 10% of Rays for Glare Enhancement

I will discuss how I then analyzed and evaluated the results.

4.1. Color Space

The raw light field data is already demosaicked into RGB color channels. The original paper suggested designating the bottom 80% of angular irradiance values as outliers and throwing them out. This left up for interpretation how that determination would be made. One method could be to independently find the outliers in each RGB channel. Another would be to treat the RGB values as 3 separate vectors and thus find the magnitudes of each combined value. Another would be to convert these RGB values to another color space altogether and sort there before converting back. Ultimately, I landed on three primary methods for outlier determination:

1. Outliers in each independent RGB channel.
2. Outliers in the luminance (Y) channel of the YCbCr color space.
3. Outliers in the lightness (L*) channel of the CIE L*a*b* color space.

It is somewhat difficult to make a qualitative comparison of the three different color spaces based solely on the final images. So, in order to look more closely at the differences, I plotted the hue and saturation values. While they aren't normalized, significant conclusions can be drawn about the effect that conversion between color spaces has on the final image quality. Refer to Figure 2 for a visual representation of the differences between the different color channels. As one can see, there is a pretty good correlation between the colors in both the original image and those manipulated in the RGB color space. Indeed, in the final images (the top row of Figure 2), the difference is barely noticeable. But a well-trained eye can see that the colors are less saturated and tend toward gray in the image manipulated in YCbCr. The difference between that manipulated in RGB and L*a*b* is even subtler still.

4.2. Percent of Light Rays Used

This section deals primarily with the number of light rays used per spatial sample and thus the number of outliers thrown away. The original Raskar et al. paper [6] suggested averaging over the lowest (aka darkest) 20% of the values in each spatial tile for glare reduction. For glare enhancement, one would simply average over the highest (aka brightest) 20%. In Figure 3, we can see clearly a comparison of between the original image (averaging over all angular values in each spatial tile) and those averaged over the darkest 10%, 25%, and 50%, respectively.

In Figure 4, we can see a quantitative difference between the luminance of the lemon image under three different conditions. The lowest luminance line (blue) corresponds to a glare reduction using only 10% of the samples in each spatial tile. The middle line (red) corresponds to the luminance of the original image (aka averaging over all of the samples in each spatial tile). And, finally, the top line (gold) corresponds to a glare enhancement where we use the brightest 10% of the samples in each spatial tiles.

4.3. Additional Statistical Methods Applied to Ray Space

For this area, I'd like to bring the reader's attention to an image of a bunch of bananas that is taken with significant background glare (Figure 5). One can see the difference between the original image (top left) and the image after the glare reduction method is applied (top right). The bunch of bananas can be clearly seen and we are also able to see more of the broom handle, window frame, and chair in the background. Since this is such a challenging photograph, I chose this one to highly a few interesting aspects that can be gleaned when you have raw light field data at your disposal. The heat maps each show some statistical operator acting on each spatial sample, with each pixel of the heat map corresponding to either the max-min differential, the average standard deviation over the RGB channels, or the median ray value at each microlens.

5. Analysis and Evaluation

5.1. Color Space

While I was able to create heat maps like those in Figure 2 for each image I analyzed and I was also able to look at min, max and mean values for each, these quantitative measures typically mirrored my own qualitative assessment. It's difficult to assign a measure of "goodness" beyond simply ensuring that you're not losing too much of the original color information. Therefore, it should come as no real surprise that manipulation in the RGB domain tended to retain the most information since that was the native color space for the raw images. However, it should be noted, that images manipulated in the YCbCr and L*a*b* color spaces did not experience significant degradation. For future analysis, however, all manipulation of ray space should just be done in the native color space unless there is a compelling reason to convert to another color space.

5.2. Percent of Light Rays Used

In Figure 3 one can clearly see that the glare off of the hood of the grill is so bright in the original image that the decal is imperceivable. In all of the subsequent images, the glare is reduced to the point where the decal is visible. One can also see that, as should be fairly logical, the fewer light rays you use, the greater chance you will not have reflection glare in those that you do use. With fewer rays, however, one also has less of an ability to manipulate the light field as one might for refocusing or other desirable image manipulations.

Looking at Figure 4, and those like it, we can get a solid quantitative comparison of the luminance in an image under different outlier conditions. While we can see a fairly large qualitative difference in the lemon images in Figure 2, the chart in Figure 4 shows just how much these values can change depending on what rays you choose in each spatial sample. Also illustrated here is the inability to easily manipulate glare in areas where the image sensor is in saturation, as in areas of significant glare (aka the plateaus in the plot).

5.3. Additional Statistical Analysis

Referring again to Figure 5, we can see clearly that we can use statistical analysis in ray space to help characterize glare in the camera and also to help devise a method for dealing with said glare. In the max-min heat map, and less so in the standard deviation heat map, we can see the hexagonal reflection glare artifacts caused by the lens elements. We can also see how these artifacts can be removed through outlier rejection, as in the background on the wall. By delving deeper into statistical methods like these, we should be able to create smarter algorithms that reduce glare while retaining more scene information that can be used to create higher resolution images.

6. Comparison to Previous Work

The results I present above are very similar to those presented in Raskar et al. [6]. It is difficult to compare one-for-one since I was unable to use the same set of sample images.

While the original paper does not go into detail on what color space was used or what specific statistical operators were examined as part of their ray space analysis (except that they averaged over the lowest 20% irradiance values for the first set of images), I believe everything I did was very much in line with what was originally done back in 2008. The only difference is that I used a commercially available, high-end, portable light field camera (the Lytro Illum).

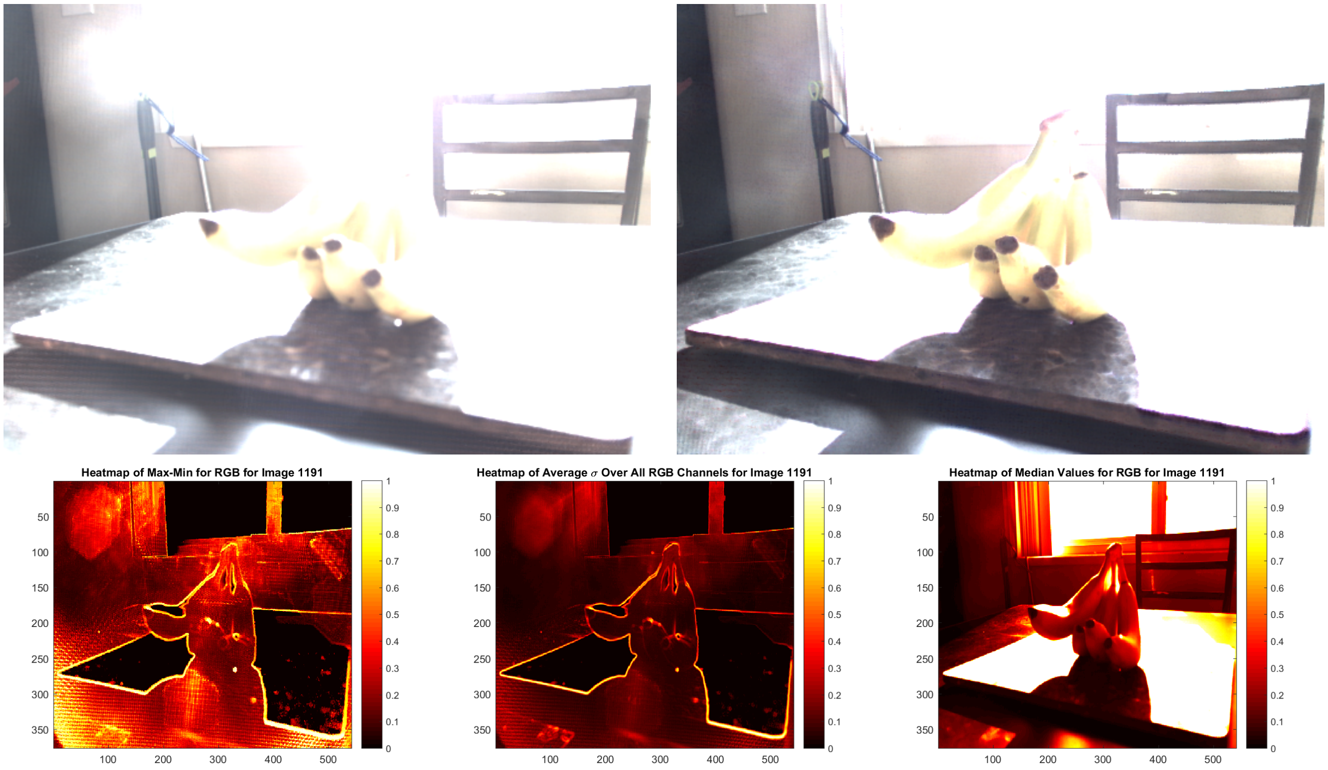

Figure 5. Statistical Analysis of An Image with Excessive Glare – [Top: Original Image (Left) and Glare Reduced Image (Right) Using RGB Manipulation Averaged over 5% of Rays / Bot: Normalized heat maps showing the difference between max and min ray values at each microlens location (Left), the standard deviation of the angular samples at each microlens location (Middle), and the average median value across the RGB channels at each microlens location (Right)]

The original paper created all of their images using a camera with pinhole array masks (both uniform and random). While fundamentally these two general approaches aren't all that different, the results were bound to differ somewhat.

Ultimately, my approach meant that I gave up significant spatial resolution over the method used by Raskar et al. However, I was able to take all of my images by hand with very short exposure times. There are also methods by which I likely could have computational reconstructed a higher resolution image using the redundant information over the angular space, but I was unable to do that in a timely and efficient manner.

Raskar et al. [6] may have been the first to significantly reduce glare using a single exposure, but they were not the last.

7. Discussion

I was able to reproduce the basic algorithm that was described in Raskar et al.'s paper [6] using a commercial, handheld light field camera (the Lytro Illum). As one can see from the images above (and the code and sample images that will be supplied with this paper), this method is very effective at doing exactly what it sets out to do. Sure, the result is a fairly low-resolution image, but you buy back image details that would normally have been washed out in a standard 2D image taken by a normal digital camera. You also get greater image contrast without losing too much of the original color information.

I originally set out with an ambition not only to reproduce the results in Raskar et al.'s original paper [6] using a commercially available light field camera, but also to incorporate and to try techniques for glare reduction that I might find elsewhere (possibly with pinhole array masks). However, my limited understanding of light fields was constrained to a lecture and a homework assignment in the class, so it took me a while to get up to speed on exactly what I was trying to do. By the time I was able to reproduce even the most basic method suggested in the paper, I had already spent considerable time. I then decided to shift course and focus on the idea of ray sampling and see what information I might be able to glean as I looked to the possibility of further characterizing glare in a 4D light field.

I attempted a couple of the other methods described in

Raskar et al.'s paper [6], but was ultimately unable to reproduce them faithfully. Specifically, I tried to implement the maxflow mincut algorithm suggested in section 4.2 I also considered attempting deconvolution by a glare point spread function (PSF), however I was unable to produce a controlled and uniform enough glare effect to have a solid chance at success. In the process of completing this project, I learned a lot about light field cameras, how one can reconstruct images using them, and their benefits/limitations. I was able to get images with significantly less glare (and much better contrast) using a simple ray sampling method. I was also able to do some basic characterization of the glare effect itself by looking at several statistical methods for evaluating ray irradiance values.

In the future, I would love to be able to use some of the statistical analysis that I developed for this project and try some more complex ways of reducing glare, such as using spatial frequency filters in ray space. I'd also like to try and get the maxflow mincut algorithm that was described in the original Raskar et al. paper [6] working as I believe it would increase performance even further. Another idea would be to try and sample the angular space of each spatial tile in a way that would reconstruct a higher resolution image from a single exposure while still reducing the glare. These are all areas that I believe I could pursue now that I have a solid understanding of the underlying concepts and especially since I have working code that can easily be manipulated to produce results quickly and efficiently.

8. Limitations

A big limitation for this particular approach is that it requires the reconstruction of a 4D light field. Even now, the resolution of microlens-based light field cameras is still pretty limited. However, with some techniques like super-resolution and angular sampling, one could conceivably reconstruct higher resolution images than those created for this project using the same camera equipment. One could also do this method using an occluded mask (as in Raskar et al.'s original paper [6], but there are drawbacks there as well (not least of which are the long exposure times). Unfortunately, this method would not be practical with a standard camera without any modification.

This method also does not work for all types of glare. It works primarily for reflection-based glare where it is limited in its angular extent. This technique would not necessarily work for diffusion-based glare as it behaves differently and relies more on the wave properties of light and doesn't respond well to the ray optics approach that this method is based upon.

Lastly, the camera used needs to have some fairly robust capabilities and be well-characterized. Using a camera like the Lytro Illum is great because each camera is well calibrated and all images take into account the individual camera's calibration data. This also helps minimize the number of artifacts created by the lenslets. The camera also needs to have a reasonable dynamic range, otherwise areas of glare may saturate the image sensor at exposures necessary to make out details in the rest of the scene. This would result in little to no useful irradiance information extracted.

9. Conclusion

As was discussed in the beginning, glare is a potential problem for many applications. Reducing or eliminating it would be incredibly beneficial under most circumstances. While the technique first proposed by Raskar et al. [6] in 2008 and reproduced for this project is not the end-all be-all, it is another step toward the better characterization and understanding of this phenomenon that has plagued photographers for generations. By looking at glare as a 4D problem and treating it in ray space, we can progress in our ability to tackle it for advanced sensors and applications. As we continue to innovate in the area of light field technology with new and improved computational imaging techniques, I believe we will only see better methods for reducing reflection and diffraction glare that take full advantage of the vast and detailed information presented to the end user.

10. Acknowledgements

I would like to thank Professor Gordon Wetzstein, Orly Liba, and Timon Ruban for a great class and a fun quarter. I personally learned a lot about computational imaging and display and am eager to learn more.

References

[1] B. Bitlis, P. A. Jansson, and J. P. Allebach. Parametric point spread function modeling and reduction of stray light effects in digital still cameras. volume 6498, pages 64980V–64980V–8, 2007.

[2] M. Levoy and P. Hanrahan. Light field rendering. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '96, pages 31–42, New York, NY, USA, 1996. ACM.

[3] K. Marwah, G. Wetzstein, A. Veeraraghavan, and R. Raskar. Compressive light field photography. In *ACM SIGGRAPH 2012 Posters*, SIGGRAPH '12, pages 49:1–49:1, New York, NY, USA, 2012. ACM.

[4] S. K. Nayar, G. Krishnan, M. D. Grossberg, and R. Raskar. Fast separation of direct and global components of a scene using high frequency illumination. In *ACM SIGGRAPH 2006 Papers*, SIGGRAPH '06, pages 935–944, New York, NY, USA, 2006. ACM.

[5] R. Ng, M. Levoy, M. Brdif, G. Duval, M. Horowitz, and P. Hanrahan. Light field photography with a hand-held plenoptic camera. Tech. rep., 2005.

[6] R. Raskar, A. Agrawal, C. A. Wilson, and A. Veeraraghavan. Glare aware photography: 4d ray sampling for reducing glare

effects of camera lenses. In *ACM SIGGRAPH 2008 Papers*, SIGGRAPH '08, pages 56:1–56:10, New York, NY, USA, 2008. ACM.

[7] E.-V. Talvala, A. Adams, M. Horowitz, and M. Levoy. Veiling glare in high dynamic range imaging. In *ACM SIGGRAPH 2007 Papers*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.

[8] G. Wetzstein, I. Ihrke, and W. Heidrich. On plenoptic multiplexing and reconstruction. *Int. J. Comput. Vision*, 101(2):384–400, Jan. 2013.